Arquitectura de software para la aplicación de técnicas de reconocimiento facial a través de dispositivos IoT.

F. Gómez, J.S. Jimenez, C.M. Blanco
Universidad Católica de Colombia
Bogotá, Colombia
fgomez17@ucatolica.edu.co
jsjimenez15@ucatolica.edu.co
cmblanco@ucatolica.edu.co

Resumen – El reconocimiento facial cada vez se vuelve más importante debido a su amplia gama de aplicaciones, pero sigue siendo un reto cuando se enfrentan a grandes variaciones en las características de los datos biométricos y sobre todo cuando se trata de transportar información a través de la red de internet en el ámbito del internet de las cosas. Este paper tiene como objetivo caracterizar los modelos de seguridad referentes al tratamiento de datos biométricos faciales, con la finalidad de proponer una arquitectura de software que contenga requerimientos de seguridad básicos necesarios, para el tratamiento de los datos involucrados en la aplicación de técnicas de reconocimiento facial, orientada a un entorno IoT.

Índice de Términos – Datos personales, reconocimiento facial, extracción de características, internet de las cosas (IoT), big Data algoritmos de clasificación, Big Data Red neuronal para aprendizaje profundo de máquina, cifrado, anti-spoofing.

I. INTRODUCCIÓN

Los equipos de cómputo y la alta tecnología han jugado un papel importante con el aforo de una mayor seguridad. Dado a su precisión, grandes bancos de memoria y alta potencia informática, se ha avanzado considerable en el área del reconocimiento facial. Los equipos de cómputo ahora sobresalen con respecto a los humanos en muchas tareas de reconocimiento facial, teniendo en cuenta que un ser humano puede recordar un número limitado de caras, pero un equipo de cómputo no tiene ningún límite y por lo tanto, puede usarse en cualquier ámbito donde sea necesario y se necesite registros faciales de autentificación. Tal es el auge del reconocimiento facial que ya se incluyó[*] en la vigilancia de multitudes, aeropuertos, seguridad privada, accesos sistemas informáticos, monitoreo de procesos automatizados. Entre otros.

[*] Andrew Meola, 'What Is the Internet of Things? IoT Definition & Meaning', Business Insider, 2018 <https://www.businessinsider.com/internet-of-things-definition> [accessed 20 May 2019]

II. ENFOQUE

A. *Autenticación*

Hace referencia al proceso electrónico mediante el cual se puede verificar la identidad de un individuo [1], buscando reducir potencialmente el fraude de identidad; esto incluye el caso en el cual un individuo intente usar sin autorización las credenciales de otro. Dentro del proceso de autenticación se han establecido tres factores que permiten realizar la identificación de quien intenta autenticarse, dichos factores involucran, *"Algo que se conoce, algo que se tiene y algo que se es"*[2].

B. *Biometría*

Conjunto de métodos automáticos que buscan identificar los individuos mediante el análisis de sus características físicas o biológicas (por ejemplo, cara, iris, huella digital, etc.). Estos métodos corresponden al tercer factor de autenticación "algo que se es", de esta forma se logra establecer medidas de control basadas en este factor, como el acceso a lugares que representan posibles riesgos en la seguridad de las organizaciones o donde se almacenen activos valiosos para las mismas [3].

C. *Internet de las cosas IoT*

Por su siglas en ingles IoT es un sistema de sensores relacionados, dispositivos informáticos y digitales repartidos por todo el mundo a través de Internet que pueden comunicarse entre ellos para compartir y transferir información utilizando una identificación única que se asigna a cada dispositivo, como UID (identificadores únicos) [4].

Estos dispositivos habitarán interconectados, con las debidas medidas de seguridad, análisis y administración; que permitan optimizar la utilización de los recursos, buscando generar conocimiento a partir de la interacción que tengan en

torno de la humanidad [5].

Utiliza diferentes tecnologías de comunicación pero no limitado a ellas com por ejemplo IPv6, Zigbee, 6LoWPAN, Bluetooth, Z-Wave, WiFi y Near Field Communications (NFC), por nombrar un pocos [56]. Tecnologías como la red centrada en la información. (ICN) y las redes definidas por software (SDN) han sido utilizado para servir como infraestructuras de comunicación subyacentes para IoT [6].

Adicionalmente se tienen una serie de características que se consideran deseables para IoT[†], a saber:

- Inteligencia Ambiental: los dispositivos deben responder al contexto de forma natural.
- Estructura Flexible: conexión y desconexión de nodos (dispositivos) de manera automática o a demanda.
- Acceso a tecnologías complejas: Esto involucra la existencia de comunicación semántica compartida entre distintos nodos de la red.

Dadas dichas aplicaciones, cabe resaltar que el presente trabajo puede ser enmarcado dentro de la categoría de vigilancia ya que presenta características de identificación de individuos.

D. Seguridad y privacidad en IoT

Las consideraciones de seguridad y privacidad de la información registrada en los dispositivos IoT no son nuevas, estas representan un desafío y es una prioridad poder garantizar que los dados que circulan en la red, solo sean accesibles para el usuario (s) que fue diseñado [7] .

Los dispositivos y servicios del IoT poco seguros pueden servir como potenciales puntos de entrada a ataques cibernéticos exponiendo los datos de los usuarios, al robo de la información por dejar flujos de datos con una protección inadecuada. La naturaleza interconectada de los dispositivos de la IoT significa que cada dispositivo mal asegurado conectado a Internet podría afectar la seguridad y la resistencia de Internet a nivel global [8].

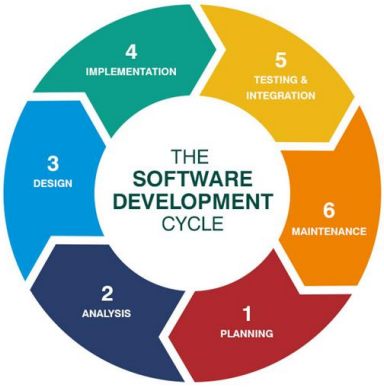
Internet de las Cosas está rediseñando los temas de privacidad de los datos registrados, dado que se pueden cambiar drásticamente las formas en que se registran, analizan, utilizan y protegen los datos personales.

Es importante considerar que, si estos retos o desafíos son significativos y muy amplios para abarcar, no son imposibles de superar.Teniendo en cuenta que la aplicación de técnicas de identificación biométrica se realiza sobre información sensible, se hace indispensable identificar requerimientos en materia de seguridad y abordarlos desde una perspectiva acorde con sus niveles de exposición a redes abiertas como el internet.

E. Aseguramiento del ciclo de vida del desarrollo de software.

Ernest Mougoue [9] describe el ciclo de vida del desarrollo de software (SDLC por sus siglas en inglés) como un marco de referencia que orienta el proceso usado en la industria para abarcar la construcción de software, desde su concepción hasta su etapa de mantenimiento.

Figura 1. Etapas del ciclo de vida de desarrollo de software.



Fuente: DATAROB[‡]

Cada una de las etapas aporta de manera diferente durante el proceso de la puesta en producción de un producto de software, este trabajo de grado se enmarca dentro de las primeras tres etapas, ya que plantea parte de los requerimientos que debería tener un software para reconocimiento facial que desee ponerse en producción sobre un entorno de IoT, el análisis ya que dichos requerimientos deben ser analizados para que sean acordes tanto a los objetivos del presente trabajo como a la realidad de la construcción de software, y diseño ya que plantea el diseño de la arquitectura de software que sea usable para enfrentar dicho problema.

Sin embargo, la seguridad de la información ha estado relegada durante años a la etapa de pruebas e integración, ya que solamente allí era donde se asignaba un espacio para las pruebas de seguridad de la información, esto conllevaba a que muchos problemas de seguridad fueran descubiertos en etapas tardías del desarrollo de un producto o que en definitiva no fueran descubiertas.

Por ello se recomienda la implementación de la seguridad de la información a través del establecimiento de controles

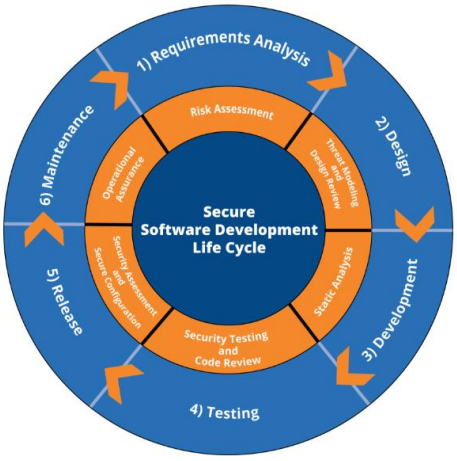
[†] jjtorres, '¿Qué Es y Cómo Funciona El Internet de Las Cosas?', 2014 <https://hipertextual.com/archivo/2014/10/internet-cosas/> [accessed 20 May 2019]

[‡] "Understanding the Software Development Life Cycle | Datarob." [Online]. Available: https://datarob.com/essentials-software-development-life-cycle/. [Accessed: 02-May-2020].

durante todo el ciclo de desarrollo de software, dando paso a un nuevo paradigma conocido como el ciclo de vida de desarrollo de software seguro (SSDLC por sus siglas en inglés). Que puede resumirse como la integración de conceptos de seguridad de la información de forma paralela durante todo el SDLC.

Entre los marcos de referencia disponibles para la aplicación del SSDLC, se encuentran los referentes de OWASP *Software Assurance Maturity Model* (SAMM)[10], *Microsoft Security development Lifecycle* (Microsoft SDL)[11], NIST SP 800-37[12], en donde cada uno de ellos se enfoca en dar recomendaciones para la implementación del SSDLC, así como las prácticas que deberían ser implementadas dentro de los controles de seguridad aplicados al SDLC para llegar a integrar niveles de seguridad avanzados al momento de gestionar el desarrollo y puesta en producción de un producto de software logrando llegar a tener un modelo de madurez como el mostrado a continuación:

Figura 2. Etapas del ciclo de vida de desarrollo seguro.



Fuente: Digital Maelstrom[13]

III. CARACTERIZACIÓN

Cada uno de los sistemas de seguridad con reconocimiento facial bajo IoT, proponen de diferentes métodos de inclusión, que conllevan a la solución de gestion de accesos automatizados. El cual están soportados en desarrollos complejos de aprendizaje profundo de máquina que conforman redes neuronales para la minería de datos, el cifrado & anti-spoofing de la información.

Dado lo anterior a continuación se relacionan algunos de los métodos de reconocimiento facial bajo dispositivos IoT.

Mehmood Irfan y otros en [14] proponen recuperar imágenes basadas en características faciales convolucionales para dispositivos con restricción de energía asistida por IoT. Además, se describe en detalle el marco de extracción de características utilizando capas intermedias de CNN[§] [15] de presentación. El sistema propuesto tiene tres pasos centrales. En primer lugar, las FN[**] caras se detectan a partir de la imagen número 1 utilizando el algoritmo Viola-Jones y se recortan solo la parte de la cara. Cada imagen puede contener un número diferente de caras. Por lo tanto, se indexaron cada cara CF[††] recortada con su imagen asociada con número 1. En segundo lugar, Las dos capas principales es decir, Conv4 y Pool3 de CNN se utilizan para la extracción de características. Las dos capas principales (es decir, Conv4 y Pool3 de CNN) se utilizan para la extracción de características. Las características de convolución (ω ^ (Conv4))[‡‡] y (ω ^ (Pool3))[§§] se fusionan como TF[***] para la representación de caras. Estos dos pasos se repiten para imágenes de conjunto de datos completos y para cada imagen las características de las caras se almacenan en la base de datos de características ΘF. Finalmente, la distancia euclidiana se utiliza para medir la puntuación de similitud entre las características de la cara de consulta Fq[†††] y la base de datos de características de la cara ΘF[‡‡‡] en tiempo real.

NETIWIT Kaongoen y otros en [16] proponen realizar un sistema de doble autenticación utilizando el modelo de clasificación P300 ERP, basados en señales cerebrales de factor único, donde se realiza los siguientes procesos: 1) entrenamiento (las fotografías de los conocidos del cliente se utilizaron como estímulos objetivo, mientras que las fotografías de personas aleatorias se seleccionaron como estímulos no objetivo); 2) autenticación propia (se genera el acceso de acuerdo al entrenamiento el cual se tiene en cuenta las dos siguientes partes); 3) ataque velado (el método de ataque en el que los atacantes conocer la información básica del objetivo, como edad, sexo, característica pero la información importante del usuario para pasar la seguridad); y 4) ataque revelado impropio (el método de ataque en el que los atacantes conocer toda la información necesaria para romper la autenticación sistema) haciendo un sistema robusto y eficiente donde se genera autoaprendizaje y se integra el gestor de seguridad.

Zhihua Xia y otros en [17] proponen un esquema para la preservación de la privacidad para el patrón binario local de la imagen en Internet industrial seguro de las cosas, donde la extracción del descriptor LBP[§§§] es fundamental ya que las imágenes se cifran mediante la combinación aleatoria de bloques y la sustitución de píxeles que preservan el orden [18] [19]. El método de encriptación de imagen diseñado de manera elaborada puede soportar el cálculo de la dirección del descriptor LBP en imágenes encriptadas donde el único que puede visualizar o modificar el usuario autorizado además que debe tener conocimiento del componente hash.

[§] Redes neuronales convolucionales.

[**] Total de caras en la imagen
[††] Cara recortada de I
[‡‡] Capa de convolución (cuatro características)
[§§] Agrupación de tres características de la capa
[***] Características fusionadas
[†††] Consulta de cara
[‡‡‡] Base de datos de características
[§§§] Local Binary Patterns (Patrones binarios locales)

IV. PROPUESTA DE LA ARQUITECTURA DE SOFTWARE.

La arquitectura cliente servidor facilita la comunicación entre diversas fuentes de información, los clientes deberán realizar cierto procesamiento con los datos, por lo cual se consideran como clientes pesados, de esta forma la capacidad de respuesta del sistema no dependerá de un servidor o un conjunto de servidores centralizados, sino que cada dispositivo de IoT que sirva el rol de un cliente el cual tendrá la responsabilidad de ejecutar diversas tareas, que contribuirán en la aplicación de algoritmos para el reconocimiento facial.

Esta propuesta, se abordó desde diferentes puntos de vista contenidos en el estándar archimate [20], estos se abordaron siguiendo una secuencia lógica de tal forma que sea comprensible la serie de pasos que llevaron al planteamiento de la presente propuesta; además se hizo especial énfasis en que esta propuesta tenga en cuenta las necesidades de protección sobre los pilares de la integridad y la confidencialidad de la información, así como las posibles etapas de desarrollo de un software que implemente esta propuesta.

A. Punto de vista de stakeholders.

Busca describir las funciones que tienen los diferentes *stakeholders* dentro del proceso de desarrollo del software, teniendo en cuenta que este desarrollo se realizaría en al menos dos fases, siendo este el punto de partida tomando lugar en las etapas de planteamiento de requerimientos, análisis y diseño, buscando incluir la seguridad de la información en las etapas tempranas, y dejando recomendaciones para las etapas de desarrollo e implementación de un prototipo funcional, se añaden los *stakeholders* que serán responsables de dicha implementación.

B. Punto de vista de función de negocio.

Se visualizan las funciones que le corresponden a los diferentes stakeholders involucrados mediante un rol específico en la realización del proyecto y sus relaciones en términos de flujo de información.

C. Punto de vista de cooperación de la aplicación.

Muestra el funcionamiento de la aplicación teniendo en cuenta la ubicación de los módulos del componente de software, sin embargo, al estar tratando un entorno de IoT se deberá plantear los diagramas correspondientes tanto a los módulos localizados en un servidor como a los ubicados en los dispositivos de IoT.

Los diagramas tienen como objetivo clarificar la comunicación entre los componentes u otras interfaces y las funciones que cumple cada uno de ellos. En este caso se tienen cuatro módulos principales, tres de ellos pertenecen al back end y la parte de interfaz de usuario que pertenece al front end. Desde el punto de vista de los dispositivos IoT se no se tendrá por defecto interfaz gráfica por lo que las respuestas de los procesos de reconocimiento facial deberán ser entregadas a través de un servicio de autenticación de usuarios el cual pueda ser consumido a través de integraciones externas.

D. Punto de vista de uso de aplicación.

Se plantea el proceso de reconocimiento facial como eje central del software, se espera que este sea capaz de desplegar diferentes funciones ofrecidas para este caso de uso.

De igual forma, se tienen los módulos base de la herramienta, cada uno de ellos presta servicios de acuerdo con su función específica. Todos estos servicios giran en torno al proceso central de verificación de identidad ya que es el eje central y punto final del desarrollo de la herramienta.

E. Punto de vista de organización e implementación.

Muestra cómo se distribuyen los módulos del prototipo dentro de la infraestructura, teniendo en cuenta la comunicación entre los diferentes, componentes lógicos y artefactos físicos.

Para este caso, tenemos componentes en infraestructura que no es propia, estos no deben quedar excluidos del diagrama y se aborda la comunicación desde las perspectivas tanto del servidor de aplicaciones, como desde los dispositivos de IoT involucrados en el proceso.

F. Punto de vista de infraestructura.

Es posible determinar la infraestructura con la que se debe contar para el desarrollo de la herramienta y así poder establecer un flujo entre los diferentes elementos involucrados.

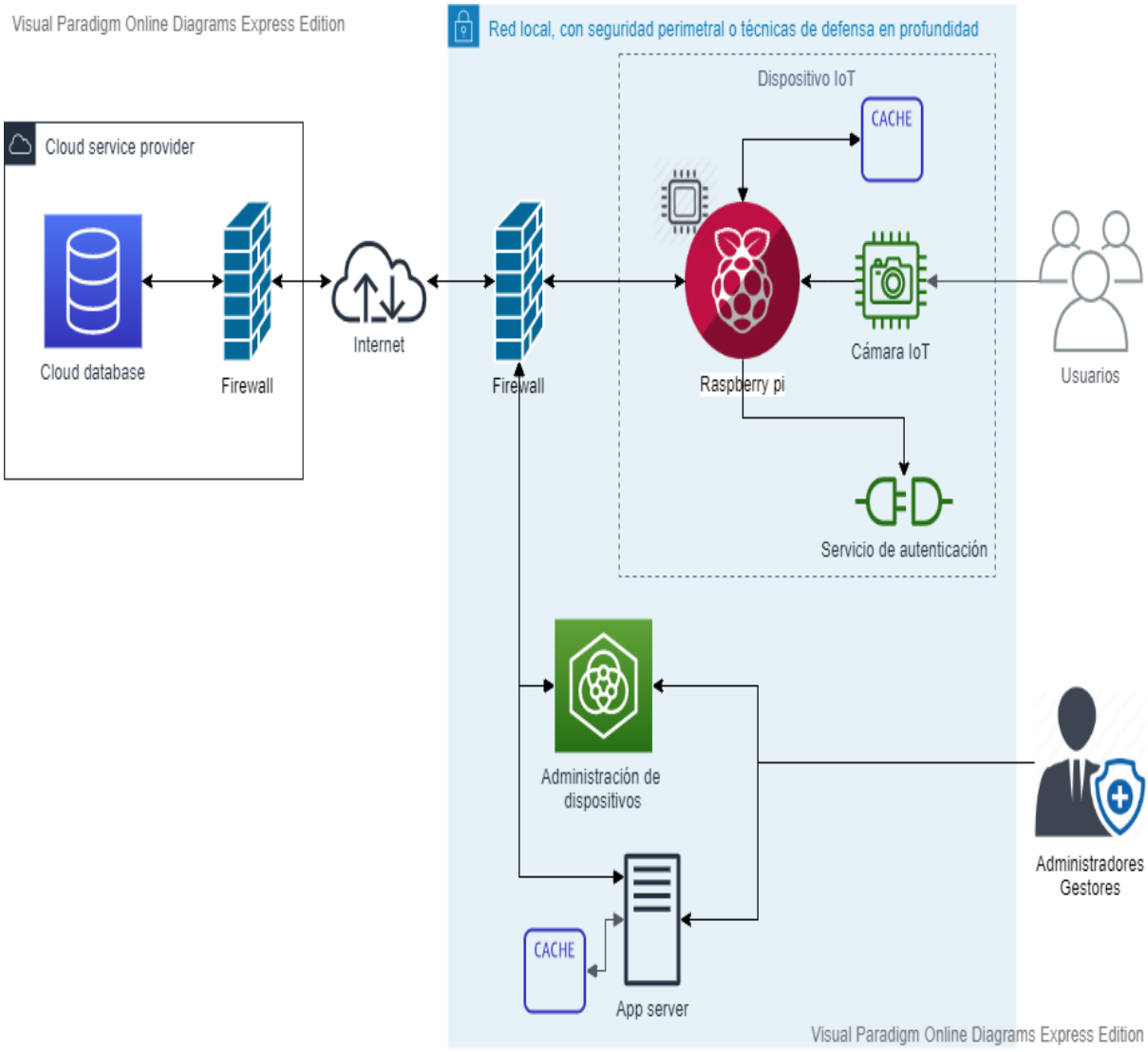
Dicha infraestructura debe ser accedida por los diferentes clientes mediante un acceso bidireccional, estableciendo así la unión entre los módulos del componente de software y el servicio que se va a prestar, se sugiere debido a la facilidad de acceso, el trabajar con una Raspberry pi como dispositivo IoT, ya que la adquisición de hardware especializado podría conllevar al aumento de costos para la realización del proyecto.

Las comunicaciones se realizarán desde la red interna en la cual se implemente el prototipo, hacia el servidor de base de datos contenido en un servicio en la nube; esto permitirá centralizar el acceso a la información en todos los dispositivos, cada uno de ellos deberá a su vez implementar una caché con lo cual reduzca la cantidad de peticiones hechas a los repositorios.

El acceso a los recursos donde se almacene la información de los usuarios debe ser restringido, y solamente se deberá transmitir información que se encuentre cifrada; ya que en

caso tal de que las comunicaciones sean interceptadas, el atacante no tendrá acceso directo a los datos que se encuentren en tránsito. Los mecanismos de cifrado serán definidos en el momento de la implementación del prototipo y deberá asegurarse que en caso de requerir alguna clave se almacene de forma confidencial y que su consulta no sea de fácil acceso, se recomienda estudiar el concepto de administración de secretos [21].

Figura 3. Diagrama de infraestructura



Fuente: los autores.

G. Nuevas Áreas De Estudio

El presente paper presentó el diseño de una arquitectura de software en la cual se abarcó parte de la etapa de planeación, análisis y diseño de un software que aplicara técnicas de reconocimiento facial en entornos de IoT integrando componentes del SSDLC; por lo tanto, se propone realizar la implementación de un prototipo funcional en el cual se evalúe la aplicación de la presente arquitectura de software, y se continúe la implementación del SSDLC en las siguientes etapas como desarrollo y pruebas.

V. CONCLUSIONES

1) En la implementación del SSDLC se requiere de la verificación exhaustiva del modelo de referencia propuesto para su uso, identificando las posibles amenazas que se podrían presentar una vez el software sea desplegado en un entorno productivo.

2) Las etapas de planeación, análisis y diseño son fundamentales para el descubrimiento de amenazas que se podrían llegar a presentar en el software, así se consigue la

implementación de controles y el planteamiento de requerimientos no funcionales que permitan disminuir la probabilidad de ocurrencia de forma temprana.

3) Los controles que se planteen a nivel de diseño requerirán del aseguramiento de la infraestructura usada mediante diversas técnicas como el aseguramiento de la red, controles de acceso y validación de identidad que permitan complementar las medidas tomadas desde el código fuente de la aplicación, así se lograría aumentar la seguridad del software en su totalidad.

4) El tratamiento de datos sensibles en entornos de IoT requiere de controles estrictos que involucran modificaciones en los procesos llevados a cabo por el software a través de las funcionalidades ofrecidas, ya que de otra forma se podrían tener brechas de seguridad que no hayan sido evaluadas.

5) Para el tratamiento de la confidencialidad de la información se requiere aplicar una serie de tratamientos a los datos de tal forma que en el caso de ser interferidos no puedan ser leídos o modificados de manera directa.

6) Por su parte el tratamiento de la integridad de la información en entornos tan complejos como el de IoT requiere que se verifique el estado de la información previo al uso de la misma, también se requiere de procesos de borrado seguro de información, mediante estos se busca eliminar cualquier rastro usable de la información en el componente de software o sus repositorios de datos.

VI. REFERENCIAS

[1] TURNER Dawn M, "Digital Authentication - the basics," 2016. .

[2] BERKELEY, "What is 3 Factor Authentication?," 2016. .

[3] MATEOS Juan, PIZARRO Sigüenza, and TAPIADOR Marino, "Tecnologías biométricas aplicadas a la seguridad," *Ra-Ma Editor.*, 2005.

[4] KUMAR Alok and JOHARI Rahul, "IOT based Electrical Device Surveillance and Control System," 2019. [Online]. Available: https://ieeexplore-ieee-org.ezproxy.javeriana.edu.co/stamp/stamp.jsp?tp=&arnumber=8777342&tag=1. [Accessed: 04-May-2020].

[5] DAVE Evans, "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything," 2011.

[6] AL-SARAWI Shadi, ANBAR Mohammed, ALIEYAN Kamal, and ALZUBAIDI Mahmood, "Internet of Things (IoT) Communication Protocols : Review," *Int. Conf. Inf. Technol.*, pp. 685–690, May 2017.

[7] SANTA RAVT Richard, "Los dispositivos IoT son un reto para la seguridad | AVI Latinoamérica," 27-Feb-2018. [Online]. Available: https://www.avilatinoamerica.com/201802275216/noticias/tecnologia/los-dispositivos-iot-son-un-reto-para-la-seguridad.html. [Accessed: 04-May-2020].

[8] ROSE Karen, ELDRIDGE Scott, and CHAPIN Lyman, "LA INTERNET DE LAS COSAS— UNA BREVE RESEÑA," Oct. 2015.

[9] MOUGOUE Ernest, "What is the secure software development life cycle (SDLC)? | Synopsys," 21-Jun-2016. [Online]. Available: https://www.synopsys.com/blogs/software-security/secure-sdlc/. [Accessed: 02-May-2020].

[10] OWASP, "OWASP SAMM," 2020. .

[11] MICROSOFT, "Microsoft Security Development Lifecycle." .

[12] TASK Joint, "SP 800-037, Rev.2, Risk Management Framework (RMF) for Information Systems and Organizations," *NIST Spec. Publ. - 800 Ser.*, p. 183, 2018.

[13] Digital Maelstrom, "Secure Software Development Lifecycle." .

[14] M. Irfan *et al.*, "Efficient Image Recognition and Retrieval on IoT-Assisted Energy-Constrained Platforms From Big Data Repositories," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9246–9255, 2019.

[15] QIANGCHANG Wang and GUODONG Guo, "LS-CNN: Characterizing Local Patches at Multiple Scales for Face Recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, no. c, pp. 1640–1653, 2020.

[16] NETIWIT Kaongoen, MOONWON Yu, and SUNGHO Jo, "Two-factor authentication system using p300 response to a sequence of human photographs," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 50, no. 3, pp. 1178–1185, 2020.

[17] ZHIHUA Xia, LEQI Jiang, XIAOHE Ma, WENYUAN Yang, PUZHAO Ji, and XIONG Neal, "A privacy-preserving outsourcing scheme for image local binary pattern in secure industrial internet of things," *IEEE Trans. Ind. Informatics*, vol. 16, no. 1, pp. 629–638, 2020.

[18] JUNG June, KIM Seung, YOO Cheol, PARK Won, and KO Sung, "LBP-ferns-based feature extraction for robust facial recognition," *IEEE Trans. Consum. Electron.*, vol. 62, no. 4, pp. 446–453, Nov. 2016.

[19] DORNAIKA Fadi and RUICHEK Yassine, "Local Binary Pattern (LBP)," *Handbook of Neural Computation*, 2017. [Online]. Available: https://www.sciencedirect.com/topics/engineering/local-binary-pattern. [Accessed: 20-May-2020].

[20] The Open Group, "The ArchiMate® Enterprise Architecture Modeling Language." .

[21] T. Reese, "Applying Secrets Management to DevOps," 2019. .